



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/822,295

DATE: 01/25/2002

TIME: 16:51:27

Input Set : N:\Crf3\RULE60\09822295.raw

Output Set: N:\CRF3\01252002\I822295.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bahija Jallal

6 Gregory D. Plowman

9 (ii) TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
10 PTP04 RELATED DISORDERS

12 (iii) NUMBER OF SEQUENCES: 18

15 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: Lyon &amp; Lyon

18 (B) STREET: 633 West Fifth Street

19 Suite 4700

20 (C) CITY: Los Angeles

21 (D) STATE: California

22 (E) COUNTRY: U.S.A.

23 (F) ZIP: 90071-2066

26 (v) COMPUTER READABLE FORM:

28 (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
29 storage

30 (B) COMPUTER: IBM Compatible

31 (C) OPERATING SYSTEM: IBM P.C. DOS 5.0

32 (D) SOFTWARE: FastSEQ for Windows 2.0

35 (vi) CURRENT APPLICATION DATA:

C--&gt; 37 (A) APPLICATION NUMBER: US/09/822,295

C--&gt; 38 (B) FILING DATE: 02-Apr-2001

39 (C) CLASSIFICATION:

42 (vii) PRIOR APPLICATION DATA:

44 (A) APPLICATION NUMBER: 09/081,345

45 (B) FILING DATE:

48 (viii) ATTORNEY/AGENT INFORMATION:

50 (A) NAME: Warburg, Richard J.

51 (B) REGISTRATION NUMBER: 32,327

52 (C) REFERENCE/DOCKET NUMBER: 234/253

55 (ix) TELECOMMUNICATION INFORMATION:

57 (A) TELEPHONE: (213) 489-1600

58 (B) TELEFAX: (213) 955-0440

59 (C) TELEX: 67-3510

64 (2) INFORMATION FOR SEQ ID NO: 1:

66 (i) SEQUENCE CHARACTERISTICS:

68 (A) LENGTH: 3580 base pairs

69 (B) TYPE: nucleic acid

70 (C) STRANDEDNESS: single

71 (D) TOPOLOGY: linear

73 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/822,295

DATE: 01/25/2002

TIME: 16:51:27

Input Set : N:\Crif3\RULE60\09822295.raw

Output Set: N:\CRF3\01252002\I822295.raw

75	CCCCGGGTGCC	CTCCCTCAAC	CTACITATAG	ACTATTTTTC	TTGCTCTGCA	GCATGGACCA	60
76	AAGAGAAATT	CTGCAGAACT	TCTCCATGA	GGCCCAAAGC	AAGAAAATTA	CTAAAGAGGA	120
77	GTTTGCCAAT	GAATTTCTGA	AGCTGAAAAG	GCAATCTACC	AAGTACAAGG	CAGACAAAAC	180
78	CTATCCTACA	ACTGTGGCTG	AGAAGCCCAA	GAATATCAAG	AAAAACAGAT	ATAAGGATAT	240
79	TTTGCCCTAT	GATTATAGCC	GGGTAGAACT	ATCCCTGATA	ACCTCTGATG	AGGATGCCAG	300
80	CTACATCAAT	GCCAACTTCA	TTAAGGGAGT	ITATGGACCC	AAGGCTTATA	TTGCCACCCA	360
81	GGGTCCCTTA	TCTACAACCC	TCTGGACTT	CTGGAGGATG	ATTTGGGAAT	ATAGTGTCTT	420
82	TATCAATTGT	ATGGCATGCA	TGGAGTATGA	AATGGGAAAG	AAAAAGTGTG	AGCGCTACTG	480
83	GGCTGAGCCA	GGAGAGATGC	AGCTGGAATT	TGGCCCTTTC	TCTGTATCCT	GTGAAGCTGA	540
84	AAAAAGGAAA	TCTGATTATA	TAATCAGGAC	TCTAAAAGTT	AAGTTCAATA	GTGAAACTCG	600
85	AACATCTTAC	CAGTTTCATT	ACAAGAATTG	GCCAGCCAT	GATGTACCTT	CATCTATAGA	660
86	CCCTATTCTT	GAGCTCATCT	GGGATGTACG	TTGTTACCAA	GAGGATGACA	GTGTTCCCAT	720
87	ATGCATTCAC	TGCAGTGCCT	GCTGTGGAAG	GACTGGTCTT	ATTTGTGCTA	TTGATTATAC	780
88	ATGGATGTTG	CTAAAAGATG	GGATAATTCC	TGAGAACTTC	AGTGTTTTCA	GTPTGATCCG	840
89	GGAAATGCGG	ACACAGAGGC	CTTCATTAGT	TCAAACGCAG	GAACAATATG	AACTGGTCTA	900
90	CAATGCTGTA	TTAGAACTAT	TTAAGAGACA	GATGGATGTT	ATCAGAGATA	AACATTCTGG	960
91	AACAGAGAGT	CAAGCAAAGC	ATTGTATTCC	TGAGAAAAAT	CACACTCTCC	AAGCAGACTC	1020
92	TTATTCTCCT	AATTTACCAA	AAAGTACCAC	AAAAGCAGCA	AAAATGATGA	ACCAACAAAAG	1080
93	GACAAAAATG	GAAATCAAAG	AATCTTCTTC	CTTTGACTTT	AGGACTTCTG	AAATAAGTGC	1140
94	AAAAGAAGAG	CTAGTTTTGC	ACCCTGCTAA	ATCAAGCACT	TCTTTTGA	TTCTGGAGCT	1200
95	AAATTACAGT	TTTGACAAAA	ATGCTGACAC	AACCATGAAA	TGGCAGACAA	AGGCATTTC	1260
96	AATAGTTGGG	GAGCCTCTTC	AGAAGCATCA	AAGTTTGGAT	TTGGGCTCTC	TTTTGTTTGA	1320
97	GGGATGTTCT	AATTCTAAAC	CTGTAATGTC	AGCAGGAAGA	TATTTTAATT	CAAAGGTGCC	1380
98	AATAACACGG	ACCAAATCAA	CTCCTTTTGA	ATTGATACAG	CAGAGAGAAA	CCAAGGAGGT	1440
99	GGACAGCAAG	GAAAACTTTT	CTTATTTGGA	ATCTCAACCA	CATGATTCTT	GTITTTGTAGA	1500
100	GAIGCAGGCT	CAAAAAGTAA	TGCAIGTTTC	TTACAGAGAA	CTGAATTATT	CACIGCCATA	1560
101	TGACTCTAAA	CACCAAATAC	GTAATGCCTC	TAATGTAAAG	CACCATGACT	CTAGTGCTCT	1620
102	TGGTGATAT	TCTTACATAC	CTTTAGTGGG	AAATCCTTAT	TTTTCATCAT	GGCCTCCAAG	1680
103	TGGTACCAGT	TCTAAGATGT	CTCTTGATTT	ACCTGAGAAG	CAAGATGGAA	CTGTTTTTCC	1740
104	TTCTTCTCTG	TTGCCAACAT	CCTCTACATC	CCTCTTCTCT	TATTACAATT	CACATGATTC	1800
105	TTTATCACTG	AATTCTCCAA	CCAATATTTT	CTCACTATTG	AACCAGGAGT	CAGCTGTACT	1860
106	AGCAACTGCI	CCAAGGATAG	ATGATGAAAT	CCCCCTCCA	CTTCTGTAC	GGACACCTGA	1920
107	ATCATTTATT	GTGGTTGAGG	AAGCTGGAGA	ATTCTCACCA	AATGTTCCCA	AATCCTTATC	1980
108	CTCAGCTGTG	AAGGTAAAAA	TTGGAACATC	ACTGGAATGG	GGTGGAACAT	CTGAACCAAA	2040
109	GAAATTTGAT	GACTCTGTGA	TACTTAGACC	AAGCAAGAGT	GTAAAACCTC	GAAGTCTTAA	2100
110	ATCAGAACTA	CATCAAGATC	GTTCTTCTCC	CCCACCTCCT	CTCCCAGAAA	GAATCTTAGA	2160
111	GTCTTCTTTT	CTTGCCGATG	AAGATTGTAT	GCAGGCCCAA	TCTATAGAAA	CATATCTTAC	2220
112	TAGCTATCCT	GACACCATGG	AAAATTCAAC	ATCTTCAAAA	CAGACACTGA	AGACTCTCTG	2280
113	AAAAAGTTTC	ACAAGGAGTA	AGAGTTTGAA	AATTTTGCGA	AACATGAAAA	AGAGTATCTG	2340
114	TAATTTCTTG	CCACCAACAA	AGCCTGCAGA	ATCTGTTTCA	TCAAATAACT	CCAGCTCATT	2400
115	TCTGAATTTT	GGTTTTGCAA	ACCGTTTTTC	AAAACCCAAA	GGACCAAGGA	ATCCACCACC	2460
116	AACITGGAAAT	ATTTAATAAA	ACTCCAGATT	TATAATAATA	TGGGCTGCAA	GTACACCTGC	2520
117	AAATAAAACT	ACTAGAATAC	TGCTAGTTAA	AATAAGTGCT	CTATATGCAT	AATATCAAAAT	2580
118	ATGAAGATAT	GCTAATGTGT	TAATAGCTTT	TAAAAGAAAA	GCAAAATGCC	AATAAGTGCC	2640
119	AGTITTTGCAT	TTTCAATATCA	TTTGCAATTGA	GTTGAAAAC	GCAAAATAAAA	GTTTGTCACT	2700
120	TGAGCTTATG	TACAGAATGC	TATATGAGAA	ACACTTTTAG	AATGGAATTA	TTTTTCATTT	2760
121	TTGCCAGTTA	TTTTTATTTT	CTTTTACTTT	TTTACATAAA	CATAAACTTC	AAAAGGTTTG	2820
122	TAAGATTITG	ATCTCAACTA	ATTTCTACAT	TGCCAGAATA	TACTATAAAA	AGTTAAAAAA	2880
123	AAACTTACTT	TGTGGGTTGC	AATACAAACT	GCTCTTGACA	ATGACTATTC	CCTGACAGTT	2940

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/822,295

DATE 01/25/2002

TIME 16:51:27

Input Set : N:\Crif3\RULE60\09822295.raw  
 Output Set: N:\CRF3\01252002\I822295.raw

```

124 ATTTTTCCTT AAATGGAGTA TACCTTGTAAT ATCTTCCCAA ATGTTGTCGA AAACITGGAAT 3000
125 ATTAAGAAAA TGAGAAATTA TATTTATTAG AATAAAATGT GCAAATAATG ACAATTATTT 3060
126 GAATGTAACA AGGAATTCAA CTGAAATCCT GATAAGTTTT AACCAAAGTC ATTAAATTAC 3120
127 CAATTCTAGA AAAGTAATCA ATGAAATATA ATAGCTATCT TTTGGTAGCA AAAGATATAA 3180
128 ATTGTATATG TTTATACAGG ATCTTTCAGA TCATGTGCAA TTTTATCTTA ACCAATCAGA 3240
129 AATACIAGTT TAAATGAAT TTCTATATGA ATATGGATCT GCCATAAGAA AATCTAGTTC 3300
130 AACTCTAATT TTATGACTA AATAAATTGG CAGGTAATTG TTTTACAAA GAATCCACCT 3360
131 GACTTCCCTT AATGCATTAA AAATATTTT ATTAAATAA CTTTATTTAT AACTTTTAGA 3420
132 AACATGTAGT ATTGTTTAAA CATCATTTGT TCTTCAGTAT TTTTCATTG GAAGTCCAAT 3480
133 AGGGCAAATT GAATGAAGTA TTATTATCIG TCTCTGTAG TACAATGTAT CCAACAGACA 3540
134 CTCAATAAAC TTTTGGTTG TTAATAAAAA AAAAAAAAAA 3580
138 (2) INFORMATION FOR SEQ ID NO: 2:
140 (i) SEQUENCE CHARACTERISTICS:
142 (A) LENGTH: 807 amino acids
143 (B) TYPE: amino acid
144 (C) STRANDEDNESS: single
145 (D) TOPOLOGY: linear
147 (ii) MOLECULE TYPE: peptide
149 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
151 Met Asp Gln Arg Glu Ile Leu Gln Lys Phe Leu Asp Glu Ala Gln Ser
152 1 5 10 15
154 Lys Lys Ile Thr Lys Glu Glu Phe Ala Asn Glu Phe Leu Lys Leu Lys
155 20 25 30
158 Arg Gln Ser Thr Lys Tyr Lys Ala Asp Lys Thr Tyr Pro Thr Thr Val
159 35 40 45
161 Ala Glu Lys Pro Lys Asn Ile Lys Lys Asn Arg Tyr Lys Asp Ile Leu
162 50 55 60
164 Pro Tyr Asp Tyr Ser Arg Val Glu Leu Ser Leu Ile Thr Ser Asp Glu
165 65 70 75 80
167 Asp Ser Ser Tyr Ile Asn Ala Asn Phe Ile Lys Gly Val Tyr Gly Pro
168 85 90 95
170 Lys Ala Tyr Ile Ala Thr Gln Gly Pro Leu Ser Thr Thr Leu Leu Asp
171 100 105 110
173 Phe Trp Arg Met Ile Trp Glu Tyr Ser Val Leu Ile Ile Val Met Ala
174 115 120 125
176 Cys Met Glu Tyr Glu Met Gly Lys Lys Lys Cys Glu Arg Tyr Trp Ala
177 130 135 140
179 Glu Pro Gly Glu Met Gln Leu Glu Phe Gly Pro Phe Ser Val Ser Cys
180 145 150 155 160
182 Glu Ala Glu Lys Arg Lys Ser Asp Tyr Ile Ile Arg Thr Leu Lys Val
183 165 170 175
185 Lys Phe Asn Ser Glu Thr Arg Thr Ile Tyr Gln Phe His Tyr Lys Asn
186 180 185 190
188 Trp Pro Asp His Asp Val Pro Ser Ser Ile Asp Pro Ile Leu Glu Leu
189 195 200 205
191 Ile Trp Asp Val Arg Cys Tyr Gln Glu Asp Asp Ser Val Pro Ile Cys
192 210 215 220
194 Ile His Cys Ser Ala Gly Cys Gly Arg Thr Gly Val Ile Cys Ala Ile
195 225 230 235 240

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/822,295

DATE: 01/25/2002

TIME: 16:51:27

Input Set : N:\Crif3\RULE60\09822295.raw

Output Set: N:\CRF3\01252002\I822295.raw

```

197 Asp Tyr Thr Trp Met Leu Leu Lys Asp Gly Ile Ile Pro Glu Asn Phe
198           245           250           255
200 Ser Val Phe Ser Leu Ile Arg Glu Met Arg Thr Gln Arg Pro Ser Leu
201           260           265           270
203 Val Gln Thr Gln Glu Gln Tyr Glu Leu Val Tyr Asn Ala Val Leu Glu
204           275           280           285
206 Leu Phe Lys Arg Gln Met Asp Val Ile Arg Asp Lys His Ser Gly Thr
207           290           295           300
210 Glu Ser Gln Ala Lys His Cys Ile Pro Glu Lys Asn His Thr Leu Gln
211 305           310           315           320
213 Ala Asp Ser Tyr Ser Pro Asn Leu Pro Lys Ser Thr Thr Lys Ala Ala
214           325           330           335
216 Lys Met Met Asn Gln Gln Arg Thr Lys Met Glu Ile Lys Glu Ser Ser
217           340           345           350
219 Ser Phe Asp Phe Arg Thr Ser Glu Ile Ser Ala Lys Glu Glu Leu Val
220           355           360           365
222 Leu His Pro Ala Lys Ser Ser Thr Ser Phe Asp Phe Leu Glu Leu Asn
223 370           375           380
225 Tyr Ser Phe Asp Lys Asn Ala Asp Thr Thr Met Lys Trp Gln Thr Lys
226 385           390           395           400
228 Ala Phe Pro Ile Val Gly Glu Pro Leu Gln Lys His Gln Ser Leu Asp
229           405           410           415
231 Leu Gly Ser Leu Leu Phe Glu Gly Cys Ser Asn Ser Lys Pro Val Asn
232           420           425           430
234 Ala Ala Gly Arg Tyr Phe Asn Ser Lys Val Pro Ile Thr Arg Thr Lys
235           435           440           445
237 Ser Thr Pro Phe Glu Leu Ile Gln Gln Arg Glu Thr Lys Glu Val Asp
238 450           455           460
240 Ser Lys Glu Asn Phe Ser Tyr Leu Glu Ser Gln Pro His Asp Ser Cys
241 465           470           475           480
243 Phe Val Glu Met Gln Ala Gln Lys Val Met His Val Ser Ser Ala Glu
244           485           490           495
246 Leu Asn Tyr Ser Leu Pro Tyr Asp Ser Lys His Gln Ile Arg Asn Ala
247           500           505           510
249 Ser Asn Val Lys His His Asp Ser Ser Ala Leu Gly Val Tyr Ser Tyr
250           515           520           525
252 Ile Pro Leu Val Glu Asn Pro Tyr Phe Ser Ser Trp Pro Pro Ser Gly
253 530           535           540
255 Thr Ser Ser Lys Met Ser Leu Asp Leu Pro Glu Lys Gln Asp Gly Thr
256 545           550           555           560
258 Val Phe Pro Ser Ser Leu Leu Pro Thr Ser Ser Thr Ser Leu Phe Ser
259           565           570           575
262 Tyr Tyr Asn Ser His Asp Ser Leu Ser Leu Asn Ser Pro Thr Asn Ile
263           580           585           590
266 Ser Ser Leu Leu Asn Gln Glu Ser Ala Val Leu Ala Thr Ala Pro Arg
267           595           600           605
269 Ile Asp Asp Glu Ile Pro Pro Pro Leu Pro Val Arg Thr Pro Glu Ser
270 610           615           620
272 Phe Ile Val Val Glu Glu Ala Gly Glu Phe Ser Pro Asn Val Pro Lys

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/822,295

DATE: 01/25/2002  
TIME: 16:51:27

Input Set : N:\Crif3\RULE60\09822295.raw  
Output Set: N:\CRF3\01252002\1822295.raw

273	625		630		635		640
275	Ser	Leu	Ser	Ser	Ala	Val	Lys
276					645		650
278	Gly	Gly	Thr	Ser	Glu	Pro	Lys
279					660		665
281	Pro	Ser	Lys	Ser	Val	Lys	Leu
282					675		680
284	Asp	Arg	Ser	Ser	Pro	Pro	Pro
285					690		695
287	Phe	Phe	Leu	Ala	Asp	Glu	Asp
288					705		710
290	Tyr	Ser	Thr	Ser	Tyr	Pro	Asp
291					725		730
293	Gln	Thr	Leu	Lys	Thr	Pro	Gly
294					740		745
296	Lys	Ile	Leu	Arg	Asn	Met	Lys
297					755		760
299	Asn	Lys	Pro	Ala	Glu	Ser	Val
300					770		775
302	Asn	Phe	Gly	Phe	Ala	Asn	Arg
303					785		790
305	Pro	Pro	Pro	Thr	Trp	Asn	Ile
306					805		

314 (2) INFORMATION FOR SEQ ID NO: 3:

- 316 (i) SEQUENCE CHARACTERISTICS:  
318 (A) LENGTH: 23 base pairs  
319 (B) TYPE: nucleic acid  
320 (C) STRANDEDNESS: single  
321 (D) TOPOLOGY: linear

323 (ix) FEATURE:

- 325 (D) OTHER INFORMATION: The letter "Y" stands for C or T.  
326 The letter "V" stands for A, C or  
327 G.

328 The letter "R" stands for A or G.

329 The letter "N" stands for A, C, G

330 or T.

332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

334 GAYTTYTGGV RNATGRNTG GGA

338 (2) INFORMATION FOR SEQ ID NO: 4:

- 340 (i) SEQUENCE CHARACTERISTICS:  
342 (A) LENGTH: 23 base pairs  
343 (B) TYPE: nucleic acid  
344 (C) STRANDEDNESS: single  
345 (D) TOPOLOGY: linear

347 (ix) FEATURE:

- 349 (D) OTHER INFORMATION: The letter "S" stands for C or G.  
350 The letter "Y" stands for C or T.  
351 The letter "N" stands for A, C, G  
352 or T.

23

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/822,295

DATE: 01/25/2002

TIME: 16:51:28

Input Set : N:\Crf3\RULE60\09822295.raw  
Output Set: N:\CRF3\01252002\I822295.raw

L:37 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:38 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:386 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:410 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6